In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [2]:
```
# carga de datos y revision rapida
data_frame = pd.read_csv('C:/Users/nuno/Downloads/Andbrain_DataSet.csv')
```

In [3]:
```
data_frame.head()
```

Out[3]:

	word	disgust	surprise	neutral	anger	sad	happy	fear
0	ability	0.004464	0.047832	0.000638	0.023597	0.013393	0.015944	0.040179
1	able	0.000017	0.000182	0.000409	0.000176	0.000219	0.000244	0.000186
2	abuse	0.000532	0.000177	0.000177	0.137363	0.001241	0.001595	0.002659
3	academy	0.007143	0.021429	0.007143	0.007143	0.007143	0.092857	0.035714
4	accept	0.008271	0.006767	0.000752	0.048872	0.018797	0.024812	0.038346

In [4]:
```
data_frame.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1104 entries, 0 to 1103
Data columns (total 8 columns):
 #   Column    Non-Null Count  Dtype  
---  ------    --------------  -----  
 0   word      1104 non-null   object 
 1   disgust   1104 non-null   float64
 2   surprise  1104 non-null   float64
 3   neutral   1104 non-null   float64
 4   anger     1104 non-null   float64
 5   sad       1104 non-null   float64
 6   happy     1104 non-null   float64
 7   fear      1104 non-null   float64
dtypes: float64(7), object(1)
memory usage: 69.1+ KB
```

In [5]:
```
# convierte las palabras en indice del data frame
data_frame = data_frame.set_index('word')
```

In [15]:
```
data_frame.mean()
```

Out[15]:
```
disgust     0.012809
surprise    0.032376
neutral     0.008170
anger       0.028224
sad         0.027243
happy       0.033701
fear        0.030917
dtype: float64
```

In [14]:
```
data_frame.median()
```

Out[14]:
```
disgust     0.007937
```

```
surprise    0.026907
neutral     0.003968
anger       0.023810
sad         0.023810
happy       0.027778
fear        0.026316
dtype: float64
```

In [13]:
```
data_frame.mode()
```

Out[13]:

	disgust	surprise	neutral	anger	sad	happy	fear
0	0.035714	0.035714	0.035714	0.035714	0.035714	0.035714	0.035714

a) De las pabras seleccionadas, la probabilidad de que tenga una connotación **feliz** es en promedio 0.3337; una media de 0.02777 el cual indica el valor a la mitad de los valores ordenados y; una moda, que indica cual es el valor más común, de 0.35714.

b) Las medidas de tendencia central son utiles para referenciar el centro de los datos. Pueden entenderse como los valores mas probables, comunes o esperados.

In [16]:
```
data_frame.var()
```

Out[16]:
```
disgust     0.000234
surprise    0.000633
neutral     0.000100
anger       0.000581
sad         0.000477
happy       0.000727
fear        0.000595
dtype: float64
```

In [17]:
```
data_frame.std()
```

Out[17]:
```
disgust     0.015313
surprise    0.025154
neutral     0.010011
anger       0.024109
sad         0.021836
happy       0.026960
fear        0.024391
dtype: float64
```

c) La varianza y desviación estandar de la probabilidad de que las pablaras tengan una connotacion **feliz** son 0.000727 y 0.026960 indican que los valores cuado es **feliz** se mueven 0.000727 respecto a su media de 0.3337.

d) Las medidas de disperción sirven para dar información de la dispersión de los datos respecto a su media.

e)

In [18]:
```
data_frame.cov()
```

Out[18]:

	disgust	surprise	neutral	anger	sad	happy	fear
disgust	0.000234	0.000057	0.000085	0.000053	0.000050	0.000021	0.000056

	disgust	surprise	neutral	anger	sad	happy	fear
surprise	0.000057	0.000633	0.000100	-0.000011	0.000057	-0.000044	0.000027
neutral	0.000085	0.000100	0.000100	0.000063	0.000081	0.000049	0.000076
anger	0.000053	-0.000011	0.000063	0.000581	0.000020	-0.000077	-0.000035
sad	0.000050	0.000057	0.000081	0.000020	0.000477	-0.000050	0.000038
happy	0.000021	-0.000044	0.000049	-0.000077	-0.000050	0.000727	-0.000070
fear	0.000056	0.000027	0.000076	-0.000035	0.000038	-0.000070	0.000595

In [19]:
```
data_frame.corr()
```

Out[19]:

	disgust	surprise	neutral	anger	sad	happy	fear
disgust	1.000000	0.149072	0.557045	0.142729	0.150864	0.049825	0.148948
surprise	0.149072	1.000000	0.396969	-0.017968	0.104156	-0.065244	0.043478
neutral	0.557045	0.396969	1.000000	0.259740	0.370187	0.181255	0.312440
anger	0.142729	-0.017968	0.259740	1.000000	0.038070	-0.118633	-0.058962
sad	0.150864	0.104156	0.370187	0.038070	1.000000	-0.085120	0.070484
happy	0.049825	-0.065244	0.181255	-0.118633	-0.085120	1.000000	-0.106338
fear	0.148948	0.043478	0.312440	-0.058962	0.070484	-0.106338	1.000000

f) las variables **anger** y **surprise**, **happy** y **suprise**, **anger** y **happy**, **anger** y **fear**, **sad** y **happy**, **happy** y **fear**